Science of the Total Environment 668 (2019) 283–294

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

# Agricultural soil organic carbon stocks in the north-eastern Iberian Peninsula: Drivers and spatial variability

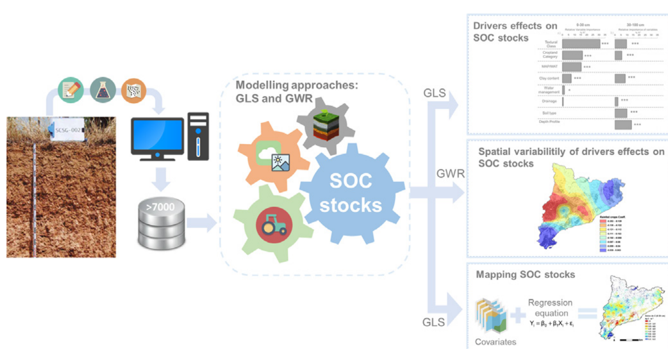
Inmaculada Funes [a,*], Robert Savé [a], Pere Rovira [c], Roberto Molowny-Horas [b], Josep M. Alcañiz [b], Emilio Ascaso [d], Ignasi Herms [d], Carmen Herrero [e], Jaume Boixadera [e], Jordi Vayreda [b]

[a] IRTA, Torre Marimon, Ctra. C-59 km 12.1, E-08140 Caldes de Montbui, Barcelona, Spain
[b] CREAF, Centre de Recerca Ecològica i Aplicacions Forestals, E-08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain
[c] CTFC, Forest Science and Technology Centre of Catalonia, Solsona 25250, Spain
[d] Àrea de Recursos Geològics, Institut Cartogràfic i Geològic de Catalunya (ICGC), E-08038 Barcelona, Spain
[e] DARP, Ministry of Agriculture, Livestock, Fisheries and Food, Government of Catalonia, Lleida E-25198, Spain

## HIGHLIGHTS

- SOC stocks were modelled using legacy data, environmental factors and geostatistics.
- The importance of SOC stock drivers differed in the top and subsoil.
- Effects of drivers on agricultural SOC stocks vary spatially at the regional scale.
- SOC stocks in Catalan agricultural soils contain $4.88 \pm 0.89$ kg/m$^2$.
- A baseline framework was established to design climate change mitigation strategies.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 12 November 2018
Received in revised form 5 February 2019
Accepted 20 February 2019
Available online 22 February 2019

Editor: Ouyang Wei

*Keywords:*
Agricultural SOC stocks
Mitigation strategies
Generalized Least Square
Geographical weighted regression
Mediterranean agriculture

## ABSTRACT

Estimating soil organic carbon (SOC) stocks under agriculture, assessing the importance of their drivers and understanding the spatial distribution of SOC stocks are crucial to predicting possible future SOC stocks scenarios under climate change conditions and to designing appropriate mitigation and adaptation strategies. This study characterized and modelled SOC stocks at two soil depth intervals, topsoil (0–30 cm) and subsoil (30–100 cm), based on both legacy and recent data from 7245 agricultural soil profiles and using environmental drivers (climate, agricultural practices and soil properties) for agricultural soils in Catalonia (NE Spain). Generalized Least Square (GLS) and Geographical Weighted Regression (GWR) were used as modelling approaches to: (i) assess the main SOC stock drivers and their effects on SOC stocks; (ii) analyse spatial variability of SOC stocks and their relationships with the main drivers; and (iii) predict and map SOC stocks at the regional scale. While topsoil variation of SOC stocks depended mainly on climate, soil texture and agricultural variables, subsoil SOC stocks changes depended mainly on soil attributes such us soil texture, clay content, soil type or depth to bedrock. The GWR model revealed that the relationship between SOC stocks and drivers varied spatially. Finally, the study was only able to predict and map topsoil SOC stocks at the regional scale, because controlling factors of SOC stocks at the subsoil level were largely unavailable for digital mapping. According to the resulting map, the mean SOC stock value for Catalan agriculture at the topsoil level was $4.88 \pm 0.89$ kg/m$^2$ and the total magnitude of the

* Corresponding author.
*E-mail address:* inmaculada.funes@irta.cat (I. Funes).

https://doi.org/10.1016/j.scitotenv.2019.02.317

carbon pool in agricultural soils of Catalonia up to 30 cm reached 47.9 Tg. The present study findings are useful for defining carbon sequestration strategies at the regional scale related with agricultural land use changes and agricultural management practices in a context of climate change.



## 1. Introduction

Soils are the largest carbon (C) terrestrial sink at global level, containing approximately 1500 Pg C at 1 m depth (Batjes, 2014), and the C stored exceeds that stored in plant biomass and the atmosphere (Vicente-Vicente et al., 2016). However, soils can become a source of atmospheric carbon dioxide ($CO_2$) depending on, for example, land use or management practices (Smith, 2012). Land use change is the leading cause of soil organic carbon (SOC) depletion at the global scale, while the main current land use change is deforestation for cultivation, particularly in subtropical and tropical countries (Canadell et al., 2007). Overall, in temperate zones, cropland areas lose >50% of their original SOC at the topsoil (0–30 cm) in about 25 to 50 years after conversion from natural ecosystems, due to changes in soil temperature, moisture regimes, soil disturbance and erosion (Lal et al., 2011). Compared to their initial status before cultivation, cropland soils covering 40–50% of global land surface (Smith, 2012) have lost about 55 Pg C worldwide (Canadell et al., 2007), but at present they still store an overall pool of 157 Pg C down to 1 m depth (Jobbagy and Jackson, 2000), which is about 10% of the global SOC pool. Fortunately, agricultural soils can be managed through the implementation of Recommended Management Practices (RMPs) to improve and restore SOC content and soil properties (Lal et al., 2011). In this respect, mitigation strategies such as the '*4 per mille Soils for Food Security and Climate*', launched at COP21, seek to increase global soil organic matter stock as a compensation for the global emissions of GHGs, in this case, by increasing SOC 0.4% per year in the first two meters of soil (Minasny et al., 2017).

Understanding the current spatial distribution of SOC stocks and its main drivers will help to predict SOC stocks changes in future climate change (CC) scenarios and define CC mitigation strategies (Yigini and Panagos, 2016). Many studies have tried to illustrate the influence of environmental drivers on soil properties as a means to understand SOC distribution based on variables such as land use, soil type, parent material, topography and climate (see review in Zhang et al., 2011). It is widely known that climate variables are important drivers of SOC stock: increasing SOC is associated with higher annual precipitation and lower temperature (Fantappie et al., 2011; Hoyle et al., 2016). Soil properties can also affect SOC stocks inasmuch as organic C is stabilized by means of physical protection or chemical mechanisms (Lawrence et al., 2015).

Mediterranean agriculture is characterized by net primary productivity regulated by limiting factors and scarce resources, such as poor water availability, soil disturbance and nutrient deficiencies (Rashid and Ryan, 2004; Torrent, 2005). Limiting net primary productivity in agriculture under Mediterranean conditions consequently reduces SOC stocks in Mediterranean agricultural soils, since C inputs, such us litter, roots or crop residues, are limited. Soil C sequestration occurs if the balance between C inputs and outputs (through emissions from respiration and mineralization) is positive and finally leads to increased SOC stocks. Several meta-analyses have been performed about SOC sequestration (C inputs > C outputs) in Mediterranean agricultural systems (Aguilera et al., 2013; Vicente-Vicente et al., 2016) with reference to land uses and RMPs. A more than likely future climate scenario in the Mediterranean region entails an increase in temperatures linked with a decrease in available soil water content that would negatively affect yields and, consequently, associated soil C inputs. However, although it is widely known that warming increases microbial activity, soil moisture could act as the main driver of soil biomes in Mediterranean environments, limiting SOC losses by microbial mineralization (Alcañiz et al., 2016). At all events, water management (irrigation or soil water harvesting and storage) is critical to the feasibility of the agricultural sector in Mediterranean regions (Montanaro et al., 2017) and the avoidance of SOC losses, since available water for crops increases biomass productivity, turnover of organic matter timing and humus formation (Lal, 2001).

Measure and prediction of SOC stocks has become a key issue in the last few decades, due to the potential impacts of climate change on them. Making accurate predictions in complex systems such as soils is a challenge, because, among other issues, data on soils is very often outdated, limited and fragmented (Chiti et al., 2012; Aksoy et al., 2016). However, there is a wide range of techniques used in predicting and mapping SOC from landscape to national or continental levels (see review in Minasny et al., 2013). Modelling based on experimental data provides opportunities to quantify the impacts of different management practices and future climate change conditions on SOC stocks (Zhang et al., 2016). The definition of a SOC stock baseline (e.g. Lugato et al., 2014a or FAO, 2018) is essential for future evaluations and, particularly in agricultural ecosystems, could contribute to assessment of the starting or ending point of the stock change that may occur after land use changes (Chiti et al., 2012) or after the establishment of certain RMPs. Moreover, mapping SOC stocks based on dynamic drivers, such as crop type, management or climate, and static drivers such as soil properties or topography would contribute to a better understanding of the spatial pattern of SOC stocks in agricultural Mediterranean soils.

To date, other SOC stock assessments have been performed for soils under agriculture in the study area (national level: Rodriguez-Martin et al., 2016 and sub-national level: Alvaro-Fuentes et al., 2011), but the present study focuses particularly on SOC stocks in agricultural soils based on a database containing data from a large number of agricultural soil profiles, with a high density of sampling points, distributed throughout the study area. Moreover, the present study provides the first assessment of agricultural SOC stocks in topsoil (first 30 cm) and subsoil (30 cm to 100 cm) while considering the main SOC drivers for Catalonia (32,108 $km^2$ NE Spain), a region that is representative of the diverse Mediterranean agricultural systems. In the present study, geostatistical techniques were applied to model SOC stocks for agricultural soils in the study area based on legacy data from 7245 agricultural soil profiles.

The main objectives of this study were: i) to assess SOC stocks at two depth intervals (top and subsoil) in soil profile; ii) to identify the main explanatory variables driving SOC stocks at the regional scale; iii) to analyse the spatial variability of relationships between SOC stocks and drivers; and iv) to map SOC stocks at the regional scale using a subset of explanatory variables (climatic, topographic and agricultural management).

## 2. Material and methods

The technical flowchart of this study is shown in Fig. 1 indicating the main steps followed in this section.

### 2.1. Study area

The study area is limited to agricultural soils in the north-eastern Iberian Peninsula (Catalonia, Fig. 2). According to SIGPAC (2016), the

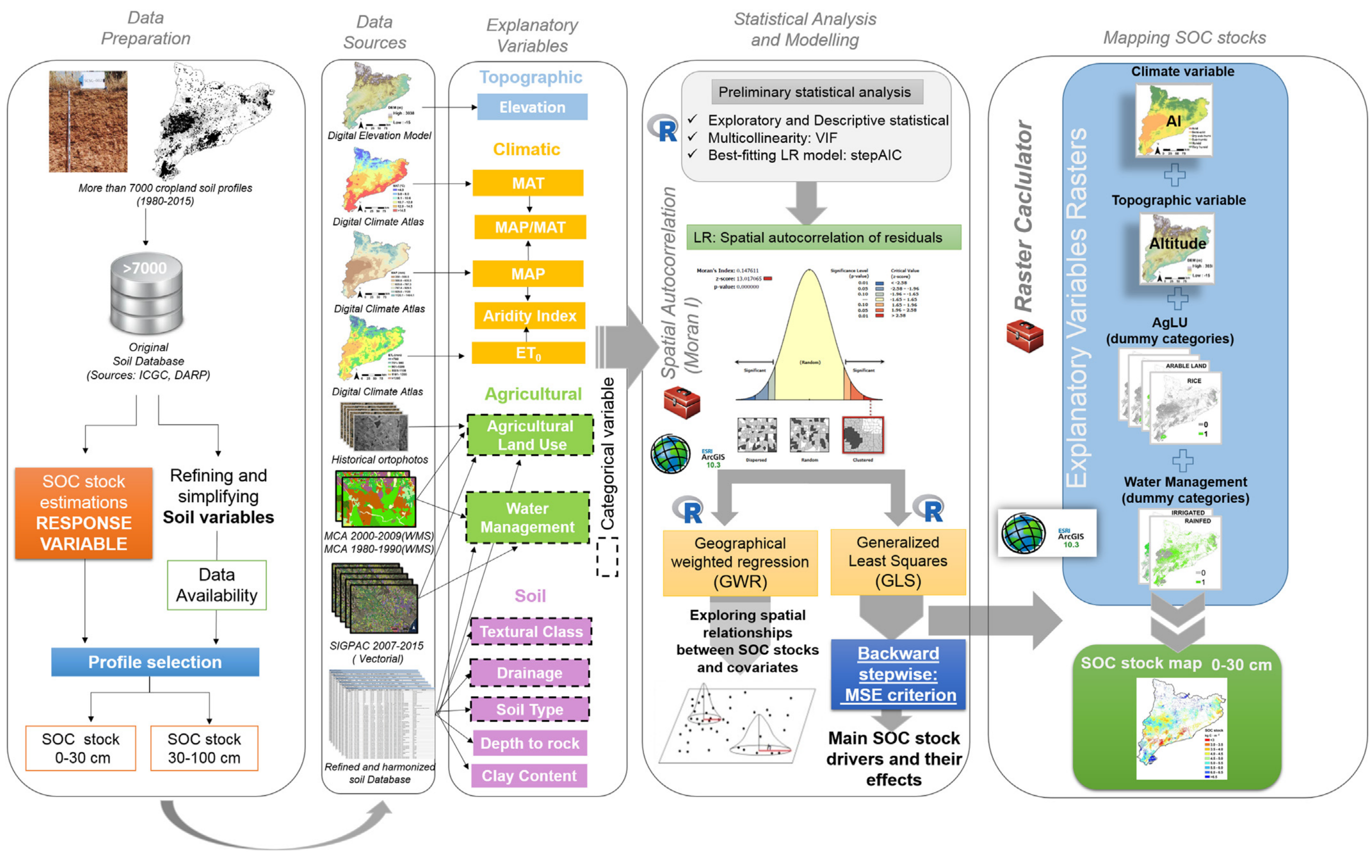



**Fig. 1.** Methodology flowchart: data, data sources, modelling framework and mapping. Climatic variables are mean annual temperature (MAT, °C), mean annual precipitation (MAP, mm), aridity index (AI, dimensionless), MAP/MAT ratio, mean annual evapotranspiration ($ET_0$, mm). Sources of both agricultural variables, Agricultural land use (AgLU) and Water Management, are: i) two editions of the map of crops and land uses (MCA, acronym in Spanish) via web map service (WMS); ii) the Agricultural Plots Geographical Information System (SIGPAC, acronym in Spanish) from the year 2007 to 2016, iii) historical ortophotos (ICGC) and iv) the original soil database. Name abbreviations in the preliminary statistical analysis: Linear Regression (LR), Mean Square Error (MSE) and Variance Inflation Factors (VIF).
(Figure inspired by Luo et al. (2017).)

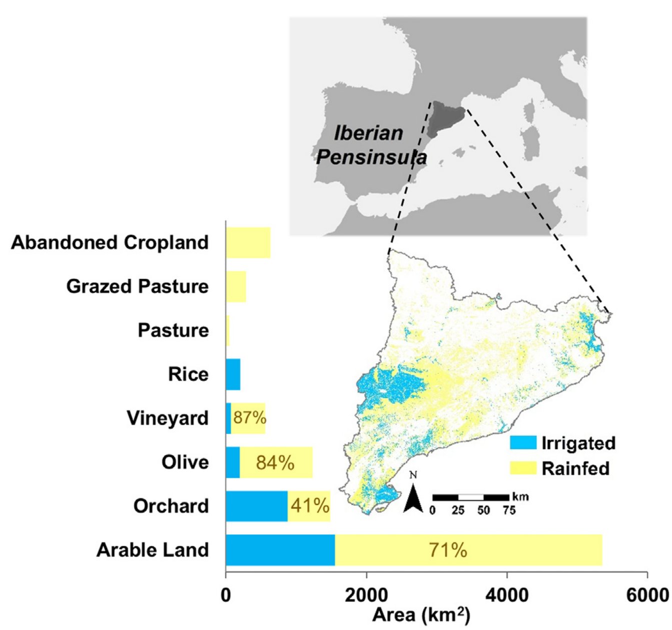



**Fig. 2.** Location of the study area. The map represents the spatial distribution of agricultural water management based on SIGPAC (2016) and DUN (2016). Area (km$^2$) per agricultural land use category is represented in the stacked bar graph; a distinction is drawn between irrigated (blue bars) and rainfed (yellow bars). Numbers within bars show the percentage of the rainfed area for each category presenting both irrigated and rainfed systems.

Agricultural Plots Geographical Information System for the year 2016, the cropland area in Catalonia is about 8837 km$^2$, not including pastures (340 km$^2$) and abandoned cropland (639 km$^2$). Almost 67% of the cropland area remains under rainfed conditions (Fig. 2). Arable land is the most widespread cropland, representing 61% of cropland area, followed by woody crops: Orchard category (17%), Olives (15%) and Vineyard (6%). Agricultural land uses extension in km$^2$ is shown in the stacked bar graph of Fig. 2, with its spatial distribution in Fig. 3. Catalonia presents Mediterranean climatic conditions characterized by mild winters and hot and dry summers (Terradas and Savé, 1992), but diverse meso- and micro-climates can be found. A strong climatic gradient (Martin-Vide et al., 2016) is defined by mean annual temperature (ranging from 0 to 17.3 °C) and annual precipitation (from 1464 mm in the Pyrenees to 335 mm in the Ebro Valley). Moreover, a marked continentality gradient is presented between inland (W) and coast (E). See more details in Fig. A.1.

Agriculture in the study area is mainly developed over Inceptisols and Entisols (mainly Fluvents and Orthents) and, to a much lesser extent, over Alfisols, Aridisols and Mollisols (SSS, 2014). Inceptisols (medium-poorly developed soils) over calcareous substrates and Entisols (very poorly developed soils) cover most of the study area. Aridisols are typical of areas where evapotranspiration is higher than precipitation, limiting crop production except when irrigation is applied, in which case high yields are obtained. Aridisols are mainly located in the Ebro valley, a historically irrigated cropland area. Agricultural soils are mostly medium (loamy) textured, with a basic reaction. Calcium carbonate-rich soils are dominant, often with a petrocalcic horizon as a root-limiting layer. Salinity problems occur over significant areas in the Ebro Valley and river deltas.

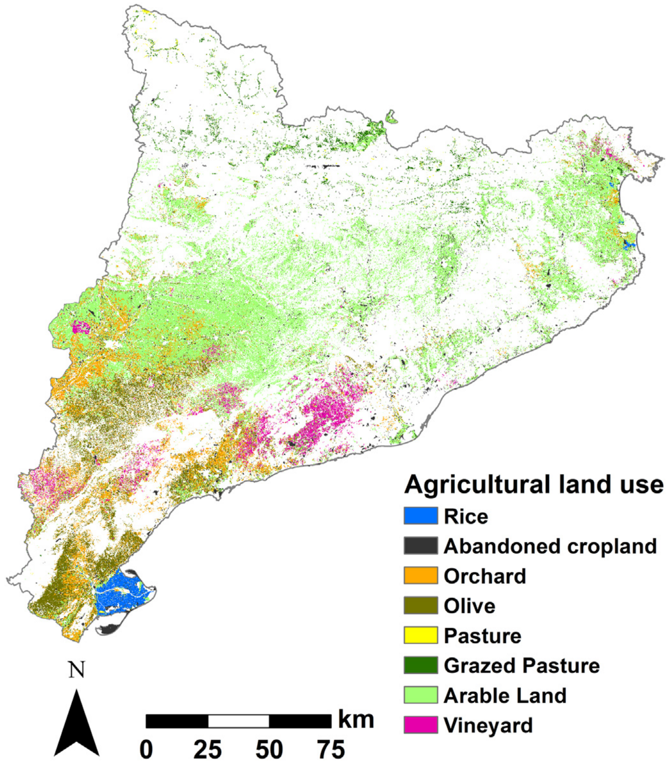



**Fig. 3.** Spatial distribution of agricultural land use categories based on SIGPAC (2016) and DUN (2016).

### 2.2. Data harmonization and SOC stock estimations

#### 2.2.1. Data collection

Soil data has been obtained from: i) the Soil database of Catalonia (BDSisCat; ICGC, 2018) of the Cartographic and Geological Institute of Catalonia (ICGC, acronym in Catalan) and ii) soil data of the Department of Agriculture, Livestock, Fisheries, Food and Environment of the Catalan Government (DARP, acronym in Catalan). Most of this data is derived from the soil survey for the Soil Map of Catalonia 1:25,000 (MSC25M; ICGC, 2017). The initial dataset included data of 7245 soil profiles, acquired from 1980 to 2015.

#### 2.2.2. SOC stock estimations

SOC stocks were estimated for two depth intervals through the vertical soil profile: standard depth intervals for topsoil (0–30 cm) and subsoil (30–100 cm). For a given horizon, the SOC stock, in $kg/m^2$, was calculated as follows:

$$\mathrm{SOC} = \mathrm{Bd} \cdot (\mathrm{OCc}/100) \cdot 10,000 \cdot \mathrm{Th} \cdot (1-\mathrm{S}) \cdot (1/1,000) \tag{1}$$

where Bd is bulk density ($g\ cm^{-3}$), OCc the concentration of OC in the fine earth (as % w/w), Th stands for the thickness of the horizon in cm, and S the stoniness (dimensionless), understood as the fraction of horizon volume (0 to 1) occupied by gravel and stones. Whereas stoniness was estimated visually in the field during soil profile description and sampling, bulk density is rarely measured in the field, and it was approached by a pedotransfer function (Honeysett and Ratkowsky, 1989). The total SOC stock of a given profile is the cumulative sum of the SOC stocks in the individual horizons, down to the desired depth.

Often this depth (either 30 or 100 cm) does not match the lowermost limit of any horizon. In such a case, it is necessary to apply a correction factor for the stock of the last horizon. Let us assume that the soil has $n$ horizons, and that the last one is divided by two by this desired depth, $d_D$. The total cumulative SOC stock, $SOC_C$, will be:

$$SOC_C = \left(\sum_{i=1}^{n-1} SOC_i\right) + \left(SOC_n \cdot \frac{d_D - d_{Un}}{d_{Ln} - d_{Un}}\right) \tag{2}$$

where $SOC_C$ is the cumulative sum of the SOC stocks of all horizons down to the desired depth (either 30 or 100 cm), $n$ stands for the horizon number which is divided by two by the desired depth, $d_D$ indicates the desired limit, $d_{Un}$ is the depth of the upper limit of this horizon, and $d_{Ln}$ is the depth of the lower limit of this horizon.

#### 2.2.3. Data selection

Profiles were excluded if data of one or more of the explanatory variables was missing (mainly soil properties; listed in Fig. 1) or if data needed to estimate SOC stocks was missing, as well. The deeper the lower limit of the depth interval, the fewer the number of profiles contained in the dataset, for usually only top horizons were analysed in soil site legacy data on organic matter and other soil properties. Spatial distribution of profiles' final dataset for each horizon (top and subsoil) is shown, respectively, in Fig. A.2.

### 2.3. Explanatory variables

A set of climatic (MAT, MAP, MAP/MAT, $ET_0$, AI; see definition of abbreviations in Fig. 1 caption), topographic (altitude), agricultural (land use and water management) and soil variables (soil texture, soil type, soil drainage, clay content and depth to bedrock) was used as potential explanatory variables for modelling SOC stocks (listed in the *Explanatory Variables* section of the methodology flowchart in Fig. 1). Detailed information about explanatory variables and their sources are explained in Appendix B of methodology.

### 2.4. Statistical analyses and modelling

Statistical and modelling analysis of the SOC stock data was conducted using the R software (R Development Core Team, 2014) and ArcGIS 10.3. (ESRI, 2011). In the analysis, the steps described below were followed:

1. Firstly, a descriptive statistics analysis (i.e. mean and standard deviation) of SOC stock data was carried out to characterize our datasets (Table 1). The results were aggregated by levels of the categorical explanatory variables.
2. Next, a preliminary visual inspection of the relationships between response and explanatory variables was performed. A careful assessment of these relationships led us to apply a square-root transformation of the SOC stock data to reduce or eliminate the impact of any heteroscedastic errors present in the data.
3. We then applied analysis-of-variance (ANOVA) to check for significant differences among the mean values of the square-root-transformed SOC stock data. To further test whether there existed significant pair-wise differences we used a post-hoc Tukey HSD test.
4. A linear regression (LR) was subsequently performed to assess the predictive power of the selected set of explanatory variables, to measure the presence of collinearity effects between them and, finally, to investigate the existence of spatial correlation in the residuals of the fit. The LR model was separately applied to the square-root-transformed top and subsoil SOC stock datasets. The starting set of explanatory variables included environmental, pedological and agricultural drivers. Once the LR model had been computed, variance inflation factors (VIFs) of the continuous explanatory variables included in the model (*vif* function in the “car” R package) were calculated to evaluate the absence of collinearity, before proceeding to eliminate those variables whose VIF was >2 (i.e. moderately to highly correlated). Next, a backward stepwise model selection strategy with all

**Table 1**
Descriptive statistics (Mean and SD) of SOC stock corresponding to 0–30 cm and 30–100 cm depth intervals, with data sets aggregated for categorical variables: agricultural land use category, water management, profile drainage, profile textural class and soil type. Relative SOC stock (%) of topsoil (0–30 cm) with respect to the first metre (0–100 cm).

| | 0–30 cm (2816 profiles) | | | 30–100 cm (1612 profiles) | | | Relative SOC stock |
|---|---|---|---|---|---|---|---|
| | n | Mean SOC (kg/m$^2$) ± SD | * | n | Mean SOC (kg/m$^2$) ± SD | * | % 0–30 cm:0–100 cm |
| Cropland category | | | | | | | |
| Grazed pasture | 141 | 6.92 ± 2.74 | a | 79 | 4.52 ± 2.81 | d | 58 |
| Rice | 49 | 6.45 ± 1.56 | a | 46 | 8.17 ± 5.15 | a | 44.3 |
| Abandoned agriculture land | 281 | 5.07 ± 1.78 | b | 140 | 6.32 ± 4.23 | b | 43.3 |
| Pasture | 38 | 4.88 ± 1.86 | bc | 25 | 4.70 ± 2.80 | cd | 51 |
| Orchard | 488 | 4.69 ± 1.59 | bc | 248 | 5.65 ± 3.01 | bc | 46.1 |
| Arable land | 1288 | 4.62 ± 1.62 | c | 808 | 6.02 ± 2.76 | b | 44.3 |
| Olive | 298 | 4.60 ± 1.37 | c | 152 | 5.13 ± 3.30 | cd | 44.8 |
| Vineyard | 233 | 3.57 ± 1.65 | d | 114 | 5.30 ± 2.61 | bcd | 42.3 |
| Water management regime | | | | | | | |
| Irrigated | 413 | 5.03 ± 1.60 | a | 262 | 6.20 ± 3.35 | a | 44.8 |
| Rainfed | 2403 | 4.68 ± 1.75 | b | 1350 | 5.75 ± 3.09 | b | 44.9 |
| Profile drainage | | | | | | | |
| Poor | 123 | 5.83 ± 1.95 | a | 86 | 7.75 ± 5.74 | a | 42.9 |
| Excessive | 117 | 4.84 ± 1.69 | b | 51 | 4.16 ± 2.93 | c | 53.8 |
| Average | 180 | 4.79 ± 1.78 | b | 129 | 5.45 ± 2.56 | b | 46.8 |
| Good | 2396 | 4.67 ± 1.70 | b | 1346 | 5.80 ± 2.90 | b | 44.6 |
| Profile textural class | | | | | | | |
| Fine | 23 | 5.88 ± 1.91 | a | 12 | 7.58 ± 2.40 | a | 43.7 |
| Medium-fine | 462 | 5.30 ± 1.79 | a | 295 | 6.93 ± 3.92 | a | 43.3 |
| Medium | 1891 | 4.72 ± 1.60 | b | 1091 | 5.75 ± 2.67 | a | 45.1 |
| Medium-coarse | 413 | 4.22 ± 1.94 | c | 198 | 4.58 ± 3.65 | b | 48.0 |
| Coarse | 27 | 2.86 ± 1.69 | d | 16 | 4.24 ± 3.29 | b | 40.3 |
| Soil type | | | | | | | |
| Mollisol | 200 | 5.65 ± 2.13 | a | 88 | 4.81 ± 3.60 | c | 54.0 |
| Aridisol | 39 | 4.77 ± 1.44 | ab | 11 | 3.08 ± 2.01 | c | 60.8 |
| Entisol | 1009 | 4.69 ± 1.65 | b | 651 | 6.50 ± 3.27 | a | 41.9 |
| Inceptisol | 1473 | 4.66 ± 1.70 | b | 808 | 5.47 ± 3.86 | b | 46.0 |
| Alfisol | 95 | 4.43 ± 1.53 | b | 54 | 5.21 ± 3.29 | bc | 46.0 |

* Same letter indicates no significant differences ($p < 0.05$) and different letter indicates differences between groups ($p < 0.05$) tested by ANOVA of square root transformed SOC stock values. SD: standard deviation, n: number of profiles.

the remaining variables for both top and subsoil datasets was performed, choosing the model with the lowest Akaike information criterion (*stepAIC* function of the "MASS" R package).

5. The residuals at every spatial location from the resulting best LR models were then determined and the corresponding Moran index was calculated (Rangel et al., 2006) using ArcGIS 10.3 for both top and subsoil datasets.
6. Once the Moran index analysis confirmed the presence of spatial correlation of residuals (For topsoil database, Moran's Index = 0.138; z-score = 12.173; $p < 0.01$ and, for subsoil database, Moran's Index = 0.095; z-score = 4.438; $p < 0.01$), their spatial correlation structure was explicitly modelled with the aid of a General Least Squares (GLS) analysis (gls function of the "nlme" R package). GLS is a regression technique by which the spatial component of the residual term is explicitly modelled in the variance-covariance matrix using parametric functions (Gaussian, exponential, lineal, etc.). In our case, GLS included X-Y site coordinates in the random-effect part of the model. Prior to the GLS calculations, data sets were partitioned into training and test subsets, containing 70% and 30% of data points, respectively. Next, backward stepwise model selection was performed, starting once again from a full model and employing the training subset for the calculations of parameter estimates. We instructed the backward stepwise procedure to remove only one or two variables at each step, due to limitations in available computing power. The criterion for model selection was mean square error (MSE), so the lower the MSE between the test data points and their corresponding predicted values (the latter determined with the parameter estimates from the model selection step), the better the model. For the spatial covariance part of the GLS model, an exponential correlation structure was chosen, which satisfactorily accounted for distance-decay effects. With the help of GLS outputs, the proportional contribution that each remaining explanatory variable made to the $R^2$ coefficient was calculated. In addition, the significance of each predictor based on the p-value of the ANOVA of best-fitting model and the relative importance of variables (RIV) in explaining variation in SOC stocks were assessed. In order to rank controlling factors of SOC stocks, the RIV of each variable was assessed as the % of the difference between the $R^2$ from the best-fitting model and the $R^2$ from the model removing each variable.
7. To further understand how the relationship between SOC stock and the explanatory variables changes spatially, a geographically weighted regression (GWR; Fotheringham and Oshan, 2016) was also performed. The GWR technique can be used to show the spatial variation of parameter estimates, which are determined locally rather than globally with a weighted least squares scheme. These weights are specified so that closer points have more influence on the determination of a local parameter than points located further away. Considering that soil samples in the present study were not regularly distributed in space, as a weighting function an adaptive spatial Gaussian kernel with a dynamically determined bandwidth was employed. GWR was performed using the same explanatory variables selected previously with the GLS model selection procedure, but this time all data points were included (i.e. without splitting the dataset into training and test subsets). Maps of continuous spatial distribution of a) GWR local estimates, b) local $R^2$ and c) residuals were generated by kriging interpolation to explore varying spatial relationships between SOC stock values and the main drivers, as well as to evaluate model performance at local and global scales.

To evaluate model performance in predicting SOC stock content, the global coefficient of determination ($R^2$), Root Mean Square Error (RMSE) and Mean Error (ME) were calculated. These indices were evaluated with the test data set for GLS models and with a complete data set for GWR.

### 2.5. Mapping SOC stocks

Digital mapping of SOC stocks was performed by applying a regression (GLS) of SOC stocks on spatial data of the environmental variables considered as predictors. Due to problems of unavailability of good spatial resolution for several covariates, such as soil properties, the covariates used in the GLS models for mapping (hereafter, GLSmap) differed from those employed in the GLS models used to assess the predictive power of each driver, as explained in step 6 of the previous section. Therefore, GLS models had to be re-fitted based on a new set of covariates. To apply the GLSmap regression equation, a set of map layers in raster format was used. Spatial data on agricultural covariates (categorical) was converted to dummy rasters (values of either 1 or 0 showing presence or absence of each category, respectively). Spatial data on climatic variables was obtained from the Digital Climatic Atlas of Catalonia and altitude data was drawn from the DEM of Catalonia (same rasters described in Appendix B of extended Material and Methods). Prediction was performed by applying map algebra (through the GLSmap regression equation and the set of covariates maps) using the GIS tool *Raster calculator* of ArcGIS 10.3.1 (ESRI, 2011). Prediction at the pixel level needed to be corrected by adding kriged residuals (differences between measured and predicted SOC stock at observed locations) from the GLS fit, in order to correct spatial correlation of residuals following Ninyerola et al. (2000). This procedure is also known as Regression Kriging in geostatistics (Chen et al., 2018).

The final corrected SOC stock maps were back-transformed from square-root in order to yield SOC stock values in kg/m$^2$ units. Spatial resolution was set to 180 × 180 m, as used in the climatic maps.

## 3. Results

### 3.1. Descriptive statistical and depth profile SOC stock distribution

Mean SOC stock values were significantly different ($p < 0.05$) for each categorical variable at the topsoil and subsoil (Table 1). Rice showed the highest mean SOC stock at the top and subsoil. Grazed pastures showed the highest mean SOC stock at the topsoil, but the lowest at the subsoil. Vineyard soils showed the lowest value in the topsoil. Irrigated cropland presented higher SOC stocks than rainfed at both top and subsoil. Poor profile drainage were associated to higher SOC stocks in both top and subsoil. With respect to textural classes, higher SOC values were linked to finer textures. Mollisols had the highest SOC stocks at the topsoil, while Entisols had the highest SOC stocks at the subsoil. Averaged values of SOC stock for agricultural land use in soil up to a depth of 1 m were ~10 kg/m$^2$, ranging from 9.2 to 14.7 kg/m$^2$, corresponding to vineyard and rice agricultural land use categories, respectively. For all categories, >50% of the total stock relative to 1 m depth was located in the subsoil (30–100 cm), except for pastures (especially grazed pasture), excessive drainages, Mollisols and Aridisols (Table 1).

### 3.2. GLS and GWR modelling, model evaluation and relative importance of explanatory variables

#### 3.2.1. GLS model performance and relative importance of variables

GLS models accounted for 27% and 20% of variations of SOC stocks at the top and subsoil, respectively (Table 2). The negative values of Mean Error (ME) obtained imply that all the prediction models were negatively unbiased, suggesting under prediction. Like $R^2$, RMSE was lower in the topsoil than in the subsoil. On the one hand, in the GLS model for topsoil significant coefficients of SOC stock (square-root transformed) were found for agricultural land use, water management and textural class. Climate variable MAP/MAT showed a positive relationship with SOC stock. However, clay content presented a significant negative coefficient of SOC stock (square-root transformed). In Fig. A.9 (a) we can see how clay content and SOC stock (square-root transformed) at the topsoil were positively correlated up to ~30 kg clay/m$^2$, and from this point on the relationship presented a slightly negative trend strongly conditioned by higher clay content, making the general trend negative. A drainage factor was not significant in the model, and soil type and soil depth were variables previously dismissed in the backward stepwise performance. On the other hand, the GLS model for subsoil showed significant coefficients for agricultural land use, excessive drainage and soil type (Table 2). Clay content and depth profiles at the subsoil showed a significant positive coefficient of SOC stock (square-root transformed). Water management and MAP/MAT were previously excluded in the backward stepwise performance.

The RIV for SOC stocks differed between top and subsoil (Fig. 4). In order to explain topsoil SOC stocks in the GLS model, textural class was the most important variable, followed by agricultural land use, MAP/MAT ratio, clay content, and finally water management. Soil properties explained 39% of topsoil SOC stock variability, land use and management 18%, and climate 15%. In the subsoil, the depth of the profile had the strongest influence on SOC stocks, followed by soil type, textural class, clay content, agricultural land use and finally drainage. Thus, soil properties explain >44% of SOC stock variability at the subsoil, whereas land use accounts for just under 6%.

#### 3.2.2. GWR model performance

GWR global coefficients of SOC stock (square-root transformed) and model evaluation (Table S2) were in line with GLS performance (Table 2). GWR coefficients of variables ranged from negative to positive, indicating the existence of spatially varying relationships between

**Table 2**

GLS model coefficients of square-root transformed SOC [sqrt SOC kg/m$^2$], confident intervals (standard error, SE), significance for each variable (based on train data set) and validation statistics (based on test data set) corresponding to 0–30 cm and 30–100 cm depth intervals.

| | | 0–30 cm GLS Model | | | 30–100 cm GLS Model | | |
|---|---|---|---|---|---|---|---|
| | | Coefficients ± SE | *t* value | | Coefficients ± SE | *t* value | |
| | Intercept | **2.821 ± 0.116** | **24.270** | *** | **1.737 ± 0.235** | **7.403** | *** |
| Respect rice | Abandoned land | **−0.176 ± 0.072** | **−2.433** | * | −0.169 ± 0.129 | −1.315 | |
| | Orchard | **−0.215 ± 0.070** | **−3.072** | ** | −0.235 ± 0.125 | −1.884 | . |
| | Olive | **−0.242 ± 0.073** | **−3.338** | *** | **−0.381 ± 0.130** | **−2.938** | ** |
| | Pasture | **−0.309 ± 0.092** | **−3.373** | *** | **−0.370 ± 0.166** | **−2.229** | * |
| | Grazed pasture | −0.147 ± 0.082 | −1.795 | . | **−0.403 ± 0.134** | **−3.006** | ** |
| | Arable Land | **−0.267 ± 0.679** | **−3.927** | *** | −0.230 ± 0.118 | −1.956 | . |
| | Vineyard | **−0.481 ± 0.737** | **−6.519** | *** | **−0.356 ± 0.134** | **−2.648** | ** |
| Respect irrigated | Rainfed | **−0.081 ± 0.024** | **−3.400** | *** | – | – | |
| Respect fine texture | Coarse | **−1.051 ± 0.112** | **−9.349** | *** | −0.279 ± 0.232 | −1.203 | |
| | Medium-fine | **−0.179 ± 0.080** | **−2.251** | * | 0.056 ± 0.174 | 0.324 | |
| | Medium-coarse | **−0.597 ± 0.087** | **−6.866** | *** | −0.334 ± 0.183 | −1.831 | . |
| | Medium | **−0.344 ± 0.082** | **−4.222** | *** | −0.121 ± 0.174 | −0.696 | |
| Respect good drainage | Poor | 0.061 ± 0.040 | 1.502 | | 0.033 ± 0.085 | 0.039 | |
| | Excessive | −0.016 ± 0.037 | −0.434 | | −0.102 ± 0.093 | −1.092 | |
| | Average | −0.044 ± 0.082 | −1.370 | | **−0.167 ± 0.062** | **−2.710** | ** |
| Respect Alfisol | Aridisol | – | – | | −0.194 ± 0.196 | −0.989 | |
| | Entisol | – | – | | **0.337 ± 0.086** | **3.912** | *** |
| | Inceptisol | – | – | | **0.171 ± 0.084** | **2.046** | * |
| | Mollisol | – | – | | 0.116 ± 0.113 | 1.027 | |
| | Depth profile | – | – | | **0.004 ± 0.0006** | **7.077** | *** |
| | Clay content | **−0.002 ± 0.0003** | **−7.176** | *** | **0.001 ± 0.0002** | **5.698** | *** |
| | MAP/MAT | **0.005 ± 0.0005** | **10.216** | *** | – | – | |
| Model evaluation (test data set) | Degrees of freedom | 722 | | | 393 | | |
| | $R^2$ | 0.272 | | | 0.203 | | |
| | ME | −0.0015 | | | −0.0011 | | |
| | RMSE | 0.318 | | | 0.573 | | |

Reference categories for each categorical predictor are denoted on the left-hand side preceded by the word *"Respect"*. *GLS Model coefficients* that are significant are shown in bold. Coefficients of variables excluded for each model in the backward stepwise performance are marked with a hyphen. An asterisk denotes $p < 0.05$, a double asterisk $p < 0.01$ and a triple asterisk $p < 0.001$. A point denotes marginal significance ($p < 0.1$).

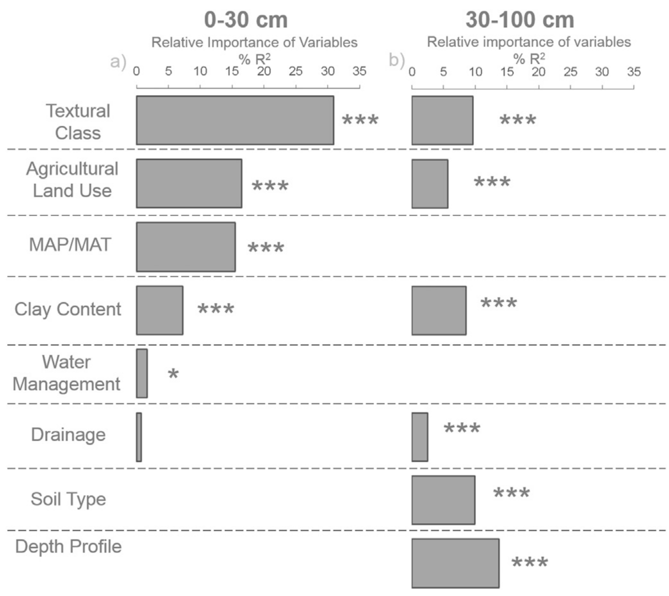



**Fig. 4.** Relative importance of variables (RIV) calculated as proportion of $R^2$ (train data set) in explaining variation of SOC stocks along a) topsoil: 0–30 cm depth interval and b) subsoil: 30–100 cm depth interval. An asterisk "*" denotes $p < 0.05$, a triple asterisk "***" denotes $p < 0.001$ and the absence of asterisks denotes no significance. Variables are sorted by relative variable importance in the GLS model set for the topsoil.

SOC stock and their explanatory variables (coefficients from the dynamic variables at the topsoil in Fig. 5 and coefficients from the rest of the explanatory variables can be found in SM). GLS and GWR global estimates for the MAP/MAT variable presented a positive sign, since the combination of high precipitation and low temperatures is related with high SOC stocks. Positive MAP/MAT coefficients at the topsoil (Fig. 5a) were distributed right across the study area, excluding the Catalan Central Depression, where negative coefficients were obtained. The obvious reason is that, in these areas, irrigation countervails drought, and high levels of plant production are attained. Although local estimates for rainfed crops (Fig.5b) were negative at the topsoil throughout the study area, the intensity of the relationships was not constant. Rainfed crops presented lower SOC stocks than irrigated right across the study area, with a more marked difference in the Catalan Central Depression. GWR coefficients for all the agricultural land use categories showed a similar spatial pattern at the topsoil (Figs. 5c and A.3 from b to g): they were negative all over Catalonia, showing the greatest magnitude in the East central areas.

The remaining variable coefficients for top and subsoil varied spatially in magnitude and even sign as well (some examples in Figs. A.3 and A.5). Higher local $R^2$ values were observed in northern areas for the topsoil GWR model (Fig. A.4a). In contrast to topsoil, higher local $R^2$ values were observed in southern zones for subsoil GWR local models (Fig. A.6a). GWR bandwidth sizes (km) were smaller for topsoil (Fig. A.4d) than subsoil (Fig. A.6d) at certain locations due to sample density (small bandwidth size was correlated to high sample density). GWR residuals were randomly spatially distributed from positive (blue colour) to negative (red colour) values at the top (Fig. A.4b) and subsoil (Fig. A.6b).

### *3.3. Mapping SOC stocks: a baseline map*

Coefficients of the GLSmap model were used at the pixel level (180 × 180 m) to predict SOC stocks (Table A.3). Explanatory variables used in the GLSmap model, limited by mapping availability and showing the best-fitting model, were: agricultural land use, water management, aridity index and altitude. Correlation coefficients of variables in the GLSmap model matched with those obtained from the GLS model used to assess the predictive power of covariates. The correlation coefficient ($R^2$) for the topsoil GLSmap model is 0.18. The agricultural soils of northern areas (Pyrenees and Pre-Pyrenees) have relatively higher SOC stocks (> 6.0 kg/m$^2$) than the rest of the region (Fig. 6). Paddy fields, found in two areas (Ebro Delta and Empordà plain), stood out with high SOC stocks. Moderate SOC stocks (4.0–5.5 kg/m$^2$) were located in the Ebro valley, southern and north-eastern regions, representing almost 84% of the study area. Soils with lower SOC stocks (<4.0 kg/m$^2$) were concentrated along the Pre-Coastal Depression (from central to south), coinciding with some important vineyard and olive growing regions. Residuals of the GLSmap model for SOC stocks at the topsoil showed spatial heterogeneity: negative (red colour) and positive (green colour) residuals, under and over predicting SOC stocks, respectively, were observed (Fig. A.7).

Averaged SOC stock values in the topsoil derived from mapping of Catalonia agriculture ranged from 0.99 to 13.98 kg/m$^2$ and the mean value was 4.88 ± 0.89 kg/m$^2$. Estimation of absolute values of SOC stocks for the topsoil total 47.89 Tg for all the agricultural land in the study area (Table 3). Most agricultural land uses (arable land, orchard, olive and abandoned land) presented a Gaussian-like distribution of SOC stock classes (i.e. symmetric histograms with most of the surface bunched in the middle SOC stock classes: from 4 to 5.5 kg/m$^2$), unlike other agricultural land uses that showed left- (rice, pastures and grazed pastures) and right- (vineyard) skewed histograms (Fig. A.8). The GLSmap model for subsoil (data not shown) indicated a negligible explained variability ($R^2 = 0.066$), and consequently mapping was dismissed.

## 4. Discussion

### *4.1. Characterizing agricultural SOC stocks and its vertical distribution up to 1 m*

The mean SOC stock values obtained from both data sets (Table 1) were in line with the previous SOC characterizations or estimations for agricultural soils down to 30 cm in other Mediterranean (Chiti et al., 2012; Rodriguez-Martin et al., 2016; Farina et al., 2017) and non-Mediterranean (Martin et al., 2011; Luo et al., 2013; Liu et al., 2015) regions.

Mean SOC stock values differed substantially from those drawn from studies in non-Mediterranean agricultural systems and other land uses. Higher values were found in agricultural soils at northern or tropical latitudes (Neufeldt, 2005; Adhikari et al., 2014; Bonfatti et al., 2016). Lower SOC stock values have been published for agricultural soils in southern, semi-arid or arid regions (Albaladejo et al., 2013; Hoyle et al., 2016; Chakan et al., 2017; Muñoz-Rojas et al., 2017; Schillaci et al., 2017a). Likewise, lower SOC stock values were estimated in Spanish soils in forest, shrubland and grassland systems estimated at 1 m depth by Doblas-Miranda et al. (2013), perhaps because these land uses are mainly encountered on shallower soils or steep slopes, whereas deeper soils and gentle slopes are preferable used for cultivated fields (Albaladejo et al., 2013; Lacoste et al., 2014) that have a greater capacity to store SOC. Notwithstanding this, when only the first 30 cm were considered, higher mean values were found under forests, shrublands and grasslands in Spain (Rodriguez-Martin et al., 2016).

The present study estimated that >50% of the total stock to 1 m depth is located in the subsoil (Table 1). These results are similar to those for soils in other climatically different regions like Iran (Chakan et al., 2017) and NW France (Lacoste et al., 2014), but are in contrast to findings for northern latitudes, where SOC stocks are greater in topsoil (Neufeldt, 2005; Kumar et al., 2013; Adhikari et al., 2014). It is commonly found that soil C generally decreases exponentially with soil depth (Albaladejo et al., 2013; Kumar et al., 2013; Hobley and Wilson, 2016).

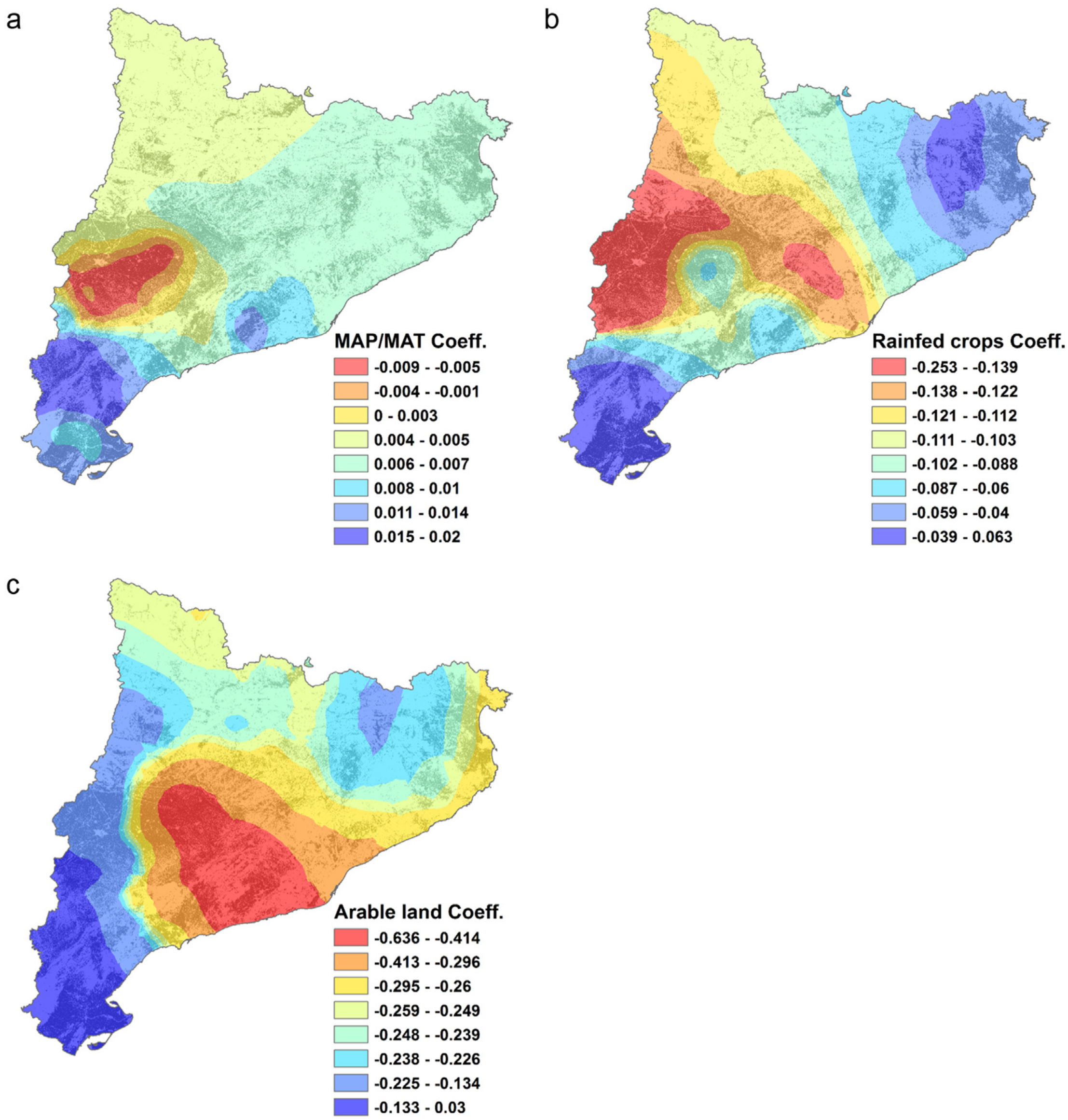



**Fig. 5.** Spatial distribution of GWR coefficients (kriged) of square-root SOC stock at 0–30 cm depth interval from explanatory variables. Maps show coefficients from: (a) MAP/MAT and (b) rainfed crops and (c) arable lands. Shadows under colours represent the spatial distribution of Catalan agriculture.

### 4.2. Modelling SOC stocks

The percentage of explained variance obtained by GLS models in this study ranged from 20% to 27%, corresponding to sub and topsoil models, respectively. Higher data density from topsoil might have a positive effect on modelling performance (Adhikari et al., 2014). $R^2$ for GLS models used to map SOC stocks in the topsoil was lower ($R^2 = 0.18$), because several drivers of SOC stock, such as soil properties, could not finally be included due to unavailability of good spatial resolution (Table A.3). In addition, for the very same reason, $R^2$ of GLS used to map SOC in the subsoil (data not shown) was negligible ($R^2 = 0.016$) and mapping SOC stocks at the subsoil was finally dismissed. Although $R^2$ values obtained may seem low, values of $R^2$ higher than 0.7 are in fact unusual, and values <0.5 are common in soil attribute prediction. Moreover, $R^2$ values usually decrease with depth (see Tables 4 and A.4; Adhikari et al., 2014; Chakan et al., 2017).

The $R^2$ values obtained could be associated with heterogeneity of spatial data density or other factors not tested due to data unavailability. Higher $R^2$ values have been found when different land uses (forest or scrubland) were modelled (Albaladejo et al., 2013).

In order to deal with spatial correlation of residuals, two models were performed: GLS and GWR. Both models are considered robust and have been widely used in statistical literature for decades (Wang et al., 2005; Rangel et al., 2006; Luo et al., 2017; Peng et al., 2017). See more references compiled in Tables 4 and A.4). Here similar results using both methodological approaches were obtained (Tables 2 and A.2).

### 4.3. Conclusive factors affecting agricultural SOC stocks

The main drivers of SOC stocks depend on the position in the soil profile (topsoil versus subsoil). At the topsoil, the main drivers were textural class, agricultural land use and MAP/MAT. Soil properties become more relevant with increasing depth. At the subsoil, the agricultural land use category was still important, but MAP/MAT ratio and water management were no longer considered important SOC drivers at depth. In line with the present study findings, some

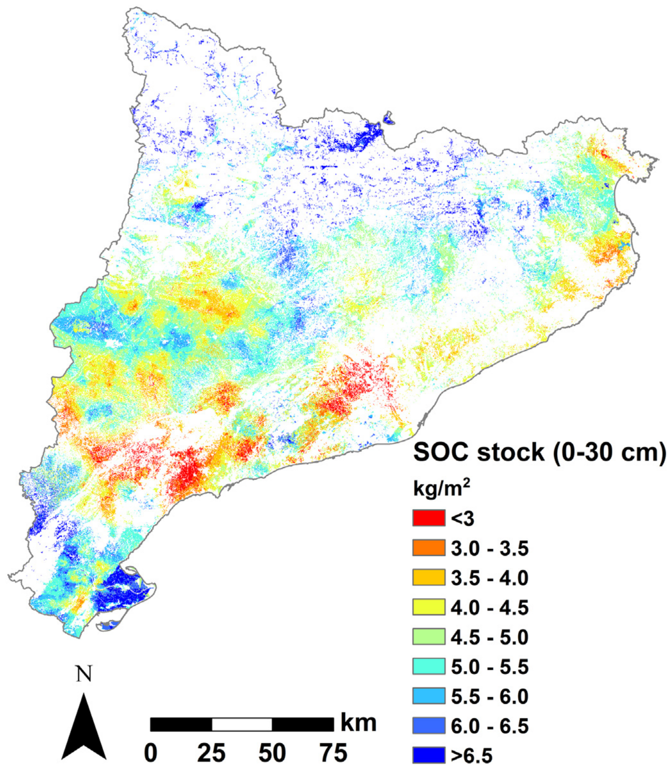



**Fig. 6.** Spatial distribution of predicted soil organic carbon stocks (kg/m$^2$) in Catalan agricultural soils using GLS performance at 0–30 cm soil depth profile. The areas rendered with white colour are non-agricultural areas.

authors (Albaladejo et al., 2013; Bonfatti et al., 2016; Armas et al., 2017; Chen et al., 2018) state that variable importance varies with depth. Climate, land use and management are likely to have a strong influence on SOC stocks at the topsoil, where these drivers directly impact. However, in the subsoil physico-chemical soil attributes are expected to be more crucial as drivers of SOC stocks than environmental factors. The importance of variables in explaining SOC stocks found here concurs with many studies (see Tables 4 and A.4) where soil properties, climate and land use and management are seen to be the key factors. Several studies have highlighted the importance of climate in predicting SOC stocks (see Tables 4 and A.4). High temperatures are related to metabolic activity stimulation of both soil microbiota and fauna, thus inducing decomposition of organic matter, while high annual precipitation relates to high net primary productivity (NPP) of plants, and hence to high inputs of organic debris to soil. C inputs are mainly limited by NPP, which depends on climate, and particularly, on the limitations in soil water and nutrients

**Table 3**
Agricultural SOC stocks of Catalonia: mean value of SOC stocks (kg/m$^2$) and absolute values of SOC pool (1 Tg = 10$^{12}$ g) aggregated by agricultural land use.

| Cropland category | Area (ha) | SOC (0–30 cm) | |
|---|---|---|---|
| | | kg/m$^2$ (mean ± SD[a]) | Tg (absolute) |
| Rice | 20,269 | 7.01 ± 0.62 | 1.42 |
| Abandoned land | 61,693 | 5.22 ± 1.24 | 3.22 |
| Orchard | 147,750 | 4.77 ± 0.94 | 7.04 |
| Olive | 123,719 | 4.70 ± 0.94 | 5.82 |
| Pastures | 5122 | 5.87 ± 1.56 | 0.30 |
| Grazed pastures | 33,933 | 6.90 ± 1.32 | 2.34 |
| Arable land | 533,651 | 4.83 ± 0.74 | 25.80 |
| Vineyard | 56,097 | 3.47 ± 0.86 | 1.95 |
| Total cropland | 982,235 | 4.88 ± 0.89 | 47.89 |

[a] SD refers to standard deviation of the predicted SOC values at the pixel level aggregated by agricultural land use category.

availability (Rabbi et al., 2015). Some studies in semi-arid Australia show that climate and soil properties better explain SOC variability compared with land use and management (Rabbi et al., 2015; Hoyle et al., 2016). Conversely, Fantappie et al. (2011) and recently Schillaci et al. (2017b) show that changes in land use and management seem to have played a major role in the variations of SOC content in Italy and Sicily, respectively.

Some authors (Jobbagy and Jackson, 2000) have pointed out that the importance of soil properties such as clay on SOC stocks increases with soil depth, playing a larger role than climate in deep layers. Given the protective role of clay, a positive impact on SOC stocks at modelling was expected. The results (Fig. A.9) show that such a positive relationship occurs only up to a given limit: about 30 kg/m$^2$ of clay in the topsoil, and about 125 kg/m$^2$ in the subsoil. From this limit on, increasing clay abundance does not result in increased SOC stocks. Indeed, a negative trend was detected in the topsoil: with very high clay stocks, SOC stocks tend to decrease. In fact, clay has a dual effect on SOC stocks (Rovira et al., 2010): positive (the protective effect on soil organic matter and the positive effect on soil water holding capacity) and negative (high amounts of clay make penetration by roots difficult, and available water for plants may be low).

### *4.4. Spatial variability of the effect of explanatory variables on SOC stocks*

The results show how at the topsoil the GWR coefficients for the climate variable MAP/MAT presented a negative counterintuitive sign in an agricultural area irrigated since the mid-19th century, the Ebro Valley (Fig. 5a). This negative relationship could be attributed to higher SOC stocks than expected in an area characterized by low precipitation and high temperatures. Possibly the impact of irrigation on the area could mask climatic effects. Rainfed coefficients were negatively stronger (Fig. 5b) at the topsoil, but only in those areas where aridity (Fig. A.1d) is more pronounced, indicating a stronger positive relationship between irrigation and SOC stocks in these semi-arid areas. Agricultural land use coefficients were negatively stronger in the middle region of Catalonia for all cropland types (Figs. 5c and A.6 from b to g), demonstrating that in this area alone SOC stocks present lower values regardless of the cropland category.

The effect of each factor on SOC stocks at top and subsoil between regions was different (Figs. 5, A.3 and A.5). Spatial variability of GWR coefficients showed how main drivers in certain locations have lesser impact, leading to a loss of importance with respect to others. This implies that when regional mitigation strategies are formulated, account should be taken of the different impact of drivers at the local scale.

### *4.5. Mapping: a new baseline*

SOC stocks were modelled and predicted assuming a steady state during the sampling period, in order that this map may be used as a baseline in the assessment of possible future spatio-temporal scenarios. Previous studies have succeeded in mapping SOC stocks at the topsoil over the study area. Using a process-based model, Alvaro-Fuentes et al. (2011) mapped SOC stocks in a wider area of NE Spain, showing values relatively far from ours in some land uses such as vineyard, olives or orchards. Notwithstanding all this, similar results have also been shown for annual and woody crops by other studies in Spain using geostatistical analysis (Rodriguez-Martin et al., 2016). The resulting SOC stocks baseline map in the present study offers improvements regarding previous baselines covering the study area. SOC stocks (topsoil) were mapped specifically for agricultural soils in Catalonia based on a high-density sampling data from >2000 spatially well-distributed agricultural soil profiles, using a statistical modelling approach and considering the main SOC drivers. Moreover, a higher map resolution for the study area was achieved, compared to existing baselines.

**Table 4**
Summary of relevant information from some of the references cited with regard to modelling SOC, important SOC drivers and agricultural SOC stock mean values.

| Region | Land uses | Model | Maximum depth (cm) | $R^2$ | SOC drivers | Cropland SOC stocks[a] (kg C/m$^2$) 0-30 cm | 0-100 cm | References |
|---|---|---|---|---|---|---|---|---|
| Catalonia, NE Spain | Cropland, grassland and unused land | GLS and GWR | 100* | 0.20-0.35 | Climate, soil properties and agricultural management | 3.57-6.92 | 9.20-14.65 | *This paper* |
| Denmark | Forest, cropland, and Wetlands | RK | 100* | 0.23-0.43 | Climate, land use, soil properties, topographic and hydrological indices. | | 12.1 | *Adhikari et al., 2014* |
| Murcia (SE Spain) | Forest, shrubland and cropland | Stepwise multiple regression analysis | 100* | 0.11-0.45 | Climate, land use, soil properties. | | 6.3 | *Albaladejo et al., 2013* |
| Western Australia | Cropland | Generalized additive mixed models | 30* | 0.72-0.79 | Climate variables, land use and agricultural management | 3.5 | | *Hoyle et al., 2016* |
| Jianghan Plain (China) | Forest, cropland, wetland and unused land | OK, MLR, GWR, RK and GWRK | 30 | 0.06-0.31 | Topography, spectral indices and distance to road | 5.03 | | *Liu et al., 2015* |
| Spain | Cropland, grassland and Forest | OK | 30 | ---- | Climate and agricultural management | 3.81-6.81 | | *Rodriguez-Martin et al., 2016* |
| France | Cropland, grassland, forest, schrubland and wetland | BRT | 30 | 0.91 | Climate, soil properties and land use | 3.2-7.57 | | *Martin et al., 2011* |
| Sicily | Cropland | SGT | 30 | 0.470 | Climate, soil properties, land use and remote sensing information | 3.5-4 | | *Schillaci et al., 2017a* |

Model name abbreviation: Ordinary Kriging (OK), Regression kriging (RK), Boosted regression tree (BRT), Multiple Linear Regression (MLR), geographically weighted regression kriging (GWRK), and Stochastic Gradient Treeboost (SGT). *analysis by depth interval.

According to Minasny et al. (2017), SOC stocks fluctuate with latitude, insofar as they are greater at higher latitudes and humid tropics and lower in the mid-latitudes. The mean SOC stock value for agriculture in the study area (4.88 kg/m$^2$) was similar to the values published for countries at similar latitudes.

### 4.6. Limitations and mapping uncertainties

In addition to the natural variability of SOC stocks, a number of different reasons could explain the low variance shown by models.

Accuracy of data for this purpose is limited. First, stoniness was estimated visually in field sampling and bulk density had to be estimated using expert-derived pedotransfer function from literature. Second, agricultural variables presented information gaps or generalizations that had to be estimated. Third, soil legacy data was sampled from 1980 to 2015, which could challenge the assumption that SOC stocks remained stable over this 35-year period, avoiding any consideration of possible climate change effects during this time. Unfortunately, due to geographical pattern of sampling, it was not possible to test the effect of sampling date on SOC stocks.

Another limitation was the lack of information related to known factors controlling SOC stocks in terms of physical or chemical C protection (Fe and Al oxides, salinity, hydromorphy, pH or clay minerals), or in terms of soil disturbance, soil protection and C inputs, such as current and historical agricultural management practices. Finally, although soil samples are well distributed right across Catalan agriculture, some agricultural areas are poorly represented. Mapping uncertainties are associated with SOC stock and driver estimations used when modelling. Modelling prediction error and unquantified uncertainties associated with some covariate layers (in some cases, rasterized versions of polygonal mapping) used to map SOC stocks should also be considered. Residuals' spatial pattern of the GLS model used for mapping SOC stocks at the topsoil (Fig. A.7) evidenced regions presenting under- and over-predictions quite consistently. These regions with higher or lower residuals (under- or over-predictions) need further attention.

### 4.7. Recommendations and future research

The results of the present study indicate that data quality must be improved to enhance modelling performance and predictions, and to reduce uncertainty in the output map. Future soil sampling efforts should focus on the acquisition of better SOC data, as well as on the collection of as many potential explanatory variables as possible (bulk density, proportion of coarse particles, detailed soil analytics, exact geographical position, detailed land use, past and current agricultural management practices, etc.). Consequently, further work must be done to understand the role of abiotic, biotic and human factors affecting spatial distribution of SOC stocks not considered here, and to build layers representing SOC stock predictors at a reasonably good spatial resolution, especially soil properties.

The resultant outputs of this study would assist in the analysis of different scenarios that help to formulate targeted climate change mitigation and adaptation policies under Mediterranean conditions. In fact, this study sets the baseline for studies exploring future climate change and land use or agriculture management scenarios, such as those published by Yigini and Panagos (2016), Lugato et al. (2014b) or Zhang et al. (2016).

## 5. Conclusions

The present study found the most important drivers of SOC stocks to be texture, climate and agricultural land use in the topsoil, and soil properties in the subsoil layer, findings that are consistent with previous studies. Topsoil offers management opportunities for C sequestration, since SOC stocks in this soil layer are mainly affected by dynamic variables. The fact that the effect of controlling factors on SOC stocks vary spatially implies that mitigation strategies should be adjusted at the local scale. Based on the available data, a modelled baseline map of SOC stocks in the topsoil (0–30 cm) for Catalan agriculture based on legacy data was produced and provided, improving spatial estimates of regional terrestrial carbon balances. Absolute and mean values of SOC stocks in soils under agriculture in Catalonia down to 30 cm are

47.89 Tg and 4.88 kg/m$^2$, respectively. This study represents a baseline framework with which to design climate change mitigation and adaptation strategies based on identifying high and low vulnerability areas and on exploring C sequestration potentials of Mediterranean agricultural soils.

## Acknowledgements

The authors would like to thank the Cartographic and Geological Institute of Catalonia and the Ministry of Agriculture, Livestock, Fisheries and Food of the Government of Catalonia for providing sampled data on soils to generate the Soil Map of Catalonia (1:25,000). The hard field work in soil sampling and subsequent analytical work undertaken by the technical personnel for the Soil Map of Catalonia during the period 1980–2015 has been critical to this study. We are particularly grateful for their important contribution to this work. This research did not receive any grant coming from the public, private, or non-profit sectors and it is the result of a collaboration agreement between our institutions (IRTA, CREAF, CTFC, ICGC and DARP). We greatly thank all people that made possible this collaboration agreement. We are grateful to both anonymous reviewers for their helpful suggestions to improve our manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2019.02.317.

## References

Adhikari, K., Hartemink, A.E., Minasny, B., Kheir, R.B., Greve, M.B., Greve, M.H., 2014. Digital mapping of soil organic carbon contents and stocks in Denmark. PLoS One 9, 13.

Aguilera, E., Lassaletta, L., Gattinger, A., Gimeno, B.S., 2013. Managing soil carbon for climate change mitigation and adaptation in Mediterranean cropping systems: a meta-analysis. Agric. Ecosyst. Environ. 168, 25–36.

Aksoy, E., Yigini, Y., Montanarella, L., 2016. Combining soil databases for topsoil organic carbon mapping in Europe. PLoS One 11, 17.

Albaladejo, J., Ortiz, R., Garcia-Franco, N., Navarro, A.R., Almagro, M., Pintado, J.G., Martínez-Mena, M., 2013. Land use and climate change impacts on soil organic carbon stocks in semi-arid Spain. J. Soils Sediments 13, 265–277.

Alcañiz, J.M., Boixadera, J., Felipó, M.T., Ortiz, J.O., Poch, R.M., 2016. Chapter 12, Sòls. In: Martin Vide, J. (Ed.), Tercer Informe sobre el Canvi Climàtic a Catalunya. Institut d'estudis Catalans i Generalitat de Catalunya, Barcelona, pp. 291–310 (614 pp.).

Alvaro-Fuentes, J., Easter, M., Cantero-Martinez, C., Paustian, K., 2011. Modelling soil organic carbon stocks and their changes in the northeast of Spain. Eur. J. Soil Sci. 62, 685–695.

Armas, D., Guevara, M., Alcaraz-Segura, D., Vargas, R., Soriano-Luna, M.A., Durante, P., Oyonarte, C., 2017. Digital map of the organic carbon profile in the soils of Andalusia, Spain. Ecosistemas 26, 80–88.

Batjes, N.H., 2014. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 65, 10–21.

Bonfatti, B.R., Hartemink, A.E., Giasson, E., Tornquist, C.G., Adhikari, K., 2016. Digital mapping of soil carbon in a viticultural region of Southern Brazil. Geoderma 261, 204–221.

Canadell, J.G., Pataki, D.E., Gifford, R., Houghton, R.A., Luo, Y., Raupach, M.R., Smith, P., Steffen, W., 2007. Saturation of the terrestrial carbon sink. In: Canadell, J.G., Pataki, D.E., Pitelka, L.F. (Eds.), Terrestrial Ecosystems in a Changing World. Springer, pp. 59–78.

Chakan, A., Taghizadeh-Mehrjardi, R., Kerry, R., Kumar, S., Khordehbin, S., Yusefi Khanghah, S., 2017. Spatial 3D distribution of soil organic carbon under different land use types. Environ. Monit. Assess. 189, 131.

Chen, S., Martin, M.P., Saby, N.P.A., Walter, C., Angers, D.A., Arrouays, D., 2018. Fine resolution map of top- and subsoil carbon sequestration potential in France. Sci. Total Environ. 630, 389–400.

Chiti, T., Gardin, L., Perugini, L., Quaratino, R., Vaccari, F.P., Miglietta, F., Valentini, R., 2012. Soil organic carbon stock assessment for the different cropland land uses in Italy. Biol. Fertil. Soils 48, 9–17.

Doblas-Miranda, E., Rovira, P., Brotons, L., Martinez-Vilalta, J., Retana, J., Pla, M., Vayreda, J., 2013. Soil carbon stocks and their variability across the forests, shrublands and grasslands of peninsular Spain. Biogeosciences 10, 8353–8361.

DUN, 2016. Agricultural Plots Geographical Information System. Shapefile. Annual. http://dadesobertes.gencat.cat.

ESRI, 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

Fantappie, M., L'Abate, G., Costantini, E.A.C., 2011. The influence of climate change on the soil organic carbon content in Italy from 1961 to 2008. Geomorphology 135, 343–352.

FAO, 2018. Soil Carbon Mapping Cookbook. Volume 2$^{nd}$ edition, Rome. ISBN 978-92-5-130440-2. http://www.fao.org/global-soil-partnership/pillars-action/4-information-and-data-new/global-soil-organic-carbon-gsoc-map/en/.

Farina, R., Marchetti, A., Francaviglia, R., Napoli, R., Bene, C.D., 2017. Modeling regional soil C stocks and CO2 emissions under Mediterranean cropping systems and soil types. Agric. Ecosyst. Environ. 238, 128–141.

Fotheringham, A.S., Oshan, T.M., 2016. Geographically weighted regression and multicollinearity: dispelling the myth. J. Geogr. Syst. 18, 303–329.

Hobley, E.U., Wilson, B., 2016. The depth distribution of organic carbon in the soils of eastern Australia. Ecosphere 7 (e01214-n/a).

Honeysett, J.L., Ratkowsky, D.A., 1989. The use of ignition to estimate bulk density of forest soils. J. Soil Sci. 40, 299–308.

Hoyle, F.C., O'Leary, R.A., Murphy, D.V., 2016. Spatially governed climate factors dominate management in determining the quantity and distribution of soil organic carbon in dryland agricultural systems. Sci. Rep. 6, 12.

ICGC, 2017. Mapa Geològic de Catalunya: Geotreball IV. Mapa de sòls. Institut Cartogràfic i Geològic de Catalunya (ICGC) http://www.icgc.cat/Ciutada/Informa-t/Recursos-geologics/Sols.

ICGC, 2018. Base de dades i Sistema d'Informació de sòls de Catalunya (BDSisCat). V1.0. Institut Cartogràfic i Geològic de Catalunya (ICGC) www.icgc.cat (CC BY 4.0).

Jobbagy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. 10, 423–436.

Kumar, S., Lal, R., Liu, D., Rafiq, R., 2013. Estimating the spatial distribution of organic carbon density for the soils of Ohio, USA. J. Geogr. Sci. 23, 280–296.

Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. Geoderma 213, 296–311.

Lal, R., 2001. World Cropland Soils as a Source or Sink for Atmospheric Carbon. Academic Press, pp. 145–191 Advances in Agronomy.

Lal, R., Delgado, J.A., Groffman, P.M., Millar, N., Dell, C., Rotz, A., 2011. Management to mitigate and adapt to climate change. J. Soil Water Conserv. 66, 276–285.

Lawrence, C.R., Harden, J.W., Xu, X., Schulz, M.S., Trumbore, S.E., 2015. Long-term controls on soil organic carbon with depth and time: a case study from the Cowlitz River Chronosequence, WA USA. Geoderma 247, 73–87.

Liu, Y., Guo, L., Jiang, Q., Zhang, H., Chen, Y., 2015. Comparing geospatial techniques to predict SOC stocks. Soil Tillage Res. 148, 46–58.

Lugato, E., Panagos, P., Bampa, F., Jones, A., Montanarella, L., 2014a. A new baseline of organic carbon stock in European agricultural soils using a modelling approach. Glob. Chang. Biol. 20, 313–326.

Lugato, E., Bampa, F., Panagos, P., Montanarella, L., Jones, A., 2014b. Potential carbon sequestration of European arable soils estimated by modelling a comprehensive set of management practices. Glob. Chang. Biol. 20, 3557–3567.

Luo, Z., Wang, E., Bryan, B.A., King, D., Zhao, G., Pan, X., Bende-Michl, U., 2013. Meta-modeling soil organic carbon sequestration potential and its application at regional scale. Ecol. Appl. 23, 408–420.

Luo, J.Q., Du, P.J., Samat, A., Xia, J.S., Che, M.Q., Xue, Z.H., 2017. Spatiotemporal pattern of PM2.5 concentrations in Mainland China and analysis of its influencing factors using geographically weighted regression. Sci. Rep. 7 (14).

Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., Arrouays, D., 2011. Spatial distribution of soil organic carbon stocks in France. Biogeosciences 8, 1053–1065.

Martin-Vide, J., Prohom Duran, M., Busto, M., 2016. Chapter 4, Evolució recent de la temperatura, la precipitació i altres variables climátiques a Catalunya. In: Martin Vide, J. (Ed.), Tercer Informe sobre el Canvi Climàtic a Catalunya. Institut d'estudis Catalans i Generalitat de Catalunya, Barcelona, pp. 93–112 (614 pp.).

Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital mapping of soil carbon. In: Sparks, D.L. (Ed.), Advances in Agronomy. vol. 118. Elsevier Academic Press Inc., San Diego, pp. 1–47.

Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L., O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G.X., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.C., Vagen, T.G., van Wesemael, B., Winowiecki, L., 2017. Soil carbon 4 per mille. Geoderma 292, 59–86.

Montanaro, G., Xiloyannis, C., Nuzzo, V., Dichio, B., 2017. Orchard management, soil organic carbon and ecosystem services in Mediterranean fruit tree crops. Sci. Hortic. 217, 92–101.

Muñoz-Rojas, M., Abd-Elmabod, S.K., Zavala, L.M., De la Rosa, D., Jordán, A., 2017. Climate change impacts on soil organic carbon stocks of Mediterranean agricultural areas: a case study in Northern Egypt. Agric. Ecosyst. Environ. 238, 142–152.

Neufeldt, H., 2005. Carbon stocks and sequestration potentials of agricultural soils in the federal state of Baden-Wurttemberg, SW Germany. J. Plant Nutr. Soil Sci.-Zeitschrift Fur Pflanzenernahrung Und Bodenkunde 168, 202–211.

Ninyerola, M., Pons, X., Roure, J.M., 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. Int. J. Climatol. 20, 1823–1841.

Peng, J., Liu, Y.X., Liu, Z.C., Yang, Y., 2017. Mapping spatial non-stationarity of human-natural factors associated with agricultural landscape multifunctionality in Beijing-Tianjin-Hebei region, China. Agric. Ecosyst. Environ. 246, 221–233.

R Development Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabbi, S.M.F., Tighe, M., Delgado-Baquerizo, M., Cowie, A., Robertson, F., Dalal, R., Page, K., Crawford, D., Wilson, B.R., Schwenke, G., McLeod, M., Badgery, W., Dang, Y.P., Bell, M., O'Leary, G., Liu, D.L., Baldock, J., 2015. Climate and soil

properties limit the positive effects of land use reversion on carbon storage in Eastern Australia. Sci. Rep. 5, 17866.
Rangel, T., Diniz-Filho, J.A.F., Bini, L.M., 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Glob. Ecol. Biogeogr. 15, 321–327.
Rashid, A., Ryan, J., 2004. Micronutrient constraints to crop production in soils with Mediterranean-type characteristics: a review. J. Plant Nutr. 27, 959–975.
Rodriguez-Martin, J.A., Alvaro-Fuentes, J., Gonzalo, J., Gil, C., Ramos-Miras, J.J., Corbi, J.M.G., Boluda, R., 2016. Assessment of the soil organic carbon stock in Spain. Geoderma 264, 117–125.
Rovira, P., Jorba, M., Romanyà, J., 2010. Active and passive organic matter fractions in Mediterranean forest soils. Biol. Fertil. Soils 46, 355–369.
Schillaci, C., Lombardo, L., Saia, S., Fantappiè, M., Märker, M., Acutis, M., 2017a. Modelling the topsoil carbon stock of agricultural lands with the Stochastic Gradient Treeboost in a semi-arid Mediterranean region. Geoderma 286, 35–45.
Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappie, M., Marker, M., Saia, S., 2017b. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: the role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. Sci. Total Environ. 601, 821–832.
SIGPAC, 2016. Crop Map of Catalonia Based on DUN. Shapefile. Annual. http://dadesobertes.gencat.cat.
Smith, P., 2012. Agricultural greenhouse gas mitigation potential globally, in Europe and in the UK: what have we learnt in the last 20 years? Glob. Chang. Biol. 18, 35–43.
SSS. Soil Survey Staff, 2014. Keys to Soil Taxonomy. twelfth edition. United States Department of Agriculture Natural Resources Conservation Service.
Terradas, J., Savé, R., 1992. Summer-winter double stress and water relationships as clue factors in the distribution of Quercus ilex L. Vegetatio 99–100, 137–145.
Torrent, J., 2005. Mediterranean soils A2 - Hillel, Daniel. Encyclopedia of Soils in the Environment. Elsevier, Oxford, pp. 418–427.
Vicente-Vicente, J.L., Garcia-Ruiz, R., Francaviglia, R., Aguilera, E., Smith, P., 2016. Soil carbon sequestration rates under Mediterranean woody crops using recommended management practices: a meta-analysis. Agric. Ecosyst. Environ. 235, 204–214.
Wang, Q., Ni, J., Tenhunen, J., 2005. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. Glob. Ecol. Biogeogr. 14, 379–393.
Yigini, Y., Panagos, P., 2016. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. Sci. Total Environ. 557, 838–850.
Zhang, C.S., Tang, Y., Xu, X.L., Kiely, G., 2011. Towards spatial geochemical modelling: use of geographically weighted regression for mapping soil organic carbon contents in Ireland. Appl. Geochem. 26, 1239–1248.
Zhang, L.M., Zhuang, Q.L., He, Y.J., Liu, Y.L., Yu, D.S., Zhao, Q.Y., Shi, X.Z., Xing, S.H., Wang, G.X., 2016. Toward optimal soil organic carbon sequestration with effects of agricultural management practices and climate change in Tai-Lake paddy soils of China. Geoderma 275, 28–39.